Targeted Gene Therapy

The present invention relates to construction of eukaryotic expression vectors and to the use of a novel methodology to regulate gene expression. More particularly, this invention relates to the use of splicing elements within such vectors to restrict desired gene translation/expression to desired environments/cell types.

With the advent of molecular biology has come the ability to express nucleotide sequences of interest (NOI) within desirable cell types. For gene therapy protocols often such expression requires targeting such that the desired NOI is expressed only within the appropriate cells. To date such targeting is mostly achieved by either 'physical' means or use of transcriptional control or both. Physical targeting can involve purification steps such that the resulting NOIs are delivered, post-purification, to the an enriched cell fraction such as CD34+ cells of blood. Other physical means by which targeting is achieved can be by way of using targeted NOI delivery vehicles (e.g. viral vectors or synthetic complexes) designed such to only deliver NOIs to the appropriate cell-type.

Transcriptional targeting often instead involves either synthetic or natural promoter elements designed to limit NOI expression to the appropriate cell types/ and or appropriate environment.

Although such targeting has great potential there are still difficulties. For physical targeting, the desired cell-types cannot always be enriched prior to NOI delivery. Furthermore targeted delivery vehicles are still being developed and consequently are often

inefficient in delivering NOI to only the desired cell type. For transcriptional targeting one problem has been that often that the synthetic or natural promoter elements, designed to confer specific NOI expression, are often 'leaky' - with transcription occurring outside of the desired environment/cell type

Achieving good cell targeting is especially desirable in the treatment of cancer. If malignant cells can be targeted and normal tissue spared then highly cyto-destructive treatment regimes can be adopted. Thus it is a particular objective of the present invention to provide a system which can target malignant cells. The NOI may itself be therapeutically active or it may be involved in a pathway which generates within the cell a therapeutically active product. The NOI may act on an endogenous but more usually an exogenous substrate to generate the therapeutically active product. In this way a prodrug (e.g. 5-fluorocytosine) may be generally administered to the patient without the need for targeting and only in the target cells expressing the NOI (e.g. cytosine deaminase) will the toxic drugs (e.g. 5-fluorouracil) be generated, typically by the enzymatic activity of the NOI.

In known techniques for targeting cancers, viral (retrovirus, adenovirus, adeno-associated virus) and non-viral (liposomes, gene gun, injection) methods have all been tried. Delivery has been targeted to tumor-specific and tissue-specific antigens and to increase specificity, tissue-specific promoters and disease-specific promoters have been used. Nevertheless, further techniques which can be used instead of or together with these methods to improve targeting are always being sought.

Presented here is a new way by which targeting can be

achieved. This invention involves the use of observed or elucidated cell-type specific splicing events to limit NOI translation to desired cell types. This targeting strategy can therefore be considered as taking place both post NOI delivery and post NOI transcription in gene therapy protocols. It however can be used in conjunction with other targeting strategies if so desired.

Although higher eukaryotes utilise promoter elements as the main method by which gene expression is restricted to desired cell types/environments, there are however other methods by when such expression can be regulated. One such method is at the level of post-transcription, and involves alternative splicing between cell types such that cell-type specific spliced transcripts are made, producing cell-type specific translated gene products.

Genes, stored as DNA sequences in the chromosomes, are transcribed into pre mRNAs which contain coding regions (exons) and generally also contain intervening non-coding regions (introns). The pre-mRNA is processed in the nucleus removing introns and any unwanted exons. The remaining exons are spliced together forming a mature mRNA molecule which is exported from the nucleus to the cytoplasm for translation into a protein by the ribosomes. It is known that different genes may be alternatively spliced in different cell types, i.e. different exons are combined to generate different proteins, see for example the Alternative Splicing Data Base at Lawrence Berkeley National Laboratory (http://devnull.lbl.gov.8888/alt). The present invention is based on the realisation that differential *cis*-splicing patterns can be used therapeutically. The particular regions, referred to herein as splice-functional elements, which direct the formation of

different splice forms have been identified in some genes and more of these regions are being characterised all the time.

One well characterised example of cell-type specific alternative splicing is that of rat B-tropomysin regulation in muscle and non-muscle cells (see Balvay *et al* 1992 NAR 20: 3987-3992; Guo and Helfman 1993 NAR 21:4762-4768). In this case, exons 1-5, 8 and 9 are common to all mRNAs expressed from this gene. However in fibroblasts and smooth muscle cells exons 6 and 11 are used, whilst in skeletal muscle cells exons 7 and 10 are exclusively used. Studies of these splicing events have localised the critical elements for alternative splicing to sequences within exon 7 and the adjacent upstream intron. Moreover these sequences have been shown to confer cell type specific splicing when flanked with heterologous exons. For example Gou and Helfman (*i.b.i.d*) have shown that the rat tropomycin exon 7 and its flanking intronic sequences are sufficient to regulate the suppression of exon 7 in non-muscle cells even when flanked by heterologous exons derived from adenovirus.

In part this invention relates to the use of such alternative splicing elements or exon/upstream intron regulatory sequences to restrict gene expression to specific cell types and/or environments.

Thus, in one aspect, the invention provides a genetic construct which comprises a nucleotide sequence of interest (NOI) and splice-functional element(s) selected and arranged such that the first open reading frame (ORF) available for translation in a target cell population is for the NOI while in non-target cells the NOI ORF is absent or is not the primary ORF.

The above defined genetic constructs include DNA molecules and pre-mRNA molecules.

The NOI open reading frame is not the primary ORF in non-target cells, (or indeed it may be absent altogether). Where it is not the primary ORF, upstream of the NOI ORF will typically be an alternative ORF which will be the primary ORF for translation. In non-target cells expression of the NOI will be non-existent or at levels less than 40%, preferably less than 30%, more preferably less than 20%, particularly preferably less than 10% of the level in target cells.

"Splice-functional elements" as defined herein are sequences present in RNA transcripts (or other corresponding DNA sequences) which are responsible for abherrant or alternative splicing.

Examples are provided herein of splice-functional elements which can be utilised to achieve cell population specific expression of a NOI and such elements which are responsible for the generation of different splice forms are being identified all the time and such sequences can be employed in novel therapies using the principles described herein. These elements are typically *cis*-acting elements found at exon and upstream intron borders. These elements may include the classical splice donor and splice acceptor sequences within exons/introns and branch points and/or sequences such as intronic distal downstream enhancers (IDDE's)which causes inclusion of an exon only in cell types which recognise the IDDE (see Example 4). Thus the splice-functional elements may result in exon inclusion or exclusion in the target cell population.

The splice-functional elements will typically but not necessarily be endogenous to the target cell-population.

Splice-functional elements from one species could be utilised in targeting a cell-population of interest from a different species, provided the required selective splice recognition had been shown.

Splice-functional elements include exons or introns (or parts thereof or flanking sequences) associated with the particular regions which are actually responsible for the cell type/environment specific splicing and are included in the construct which is introduced into the cells. These regions are typically endogenous to the target cells but are not translated in the target cell-population either independently or as fusion proteins with the product of the NOI. Regions from the splice-functional elements may be translated in non-target cells as the primary ORF which effectively prevents translation of the NOI which may be present in the mature mRNA product.

Although such sequences which direct differential splicing patterns are known, it has not previously been thought to use these in therapy to express a therapeutic gene of interest. In particular, according to the techniques described herein, the splice-functional elements are used to generate different mature mRNA molecules in target and non-target cell populations; preferably exons from the splice-functional elements do not form part of the translated protein product encoded by the NOI.

Thus in a further aspect, the invention provide the use of splice-functional elements to restrict translation or expression of a nucleotide sequence (NOI) of interest to one or more specific cell types and/or environments.

Splice-functional elements for use in the invention may be chosen from any system in which abherrant/alternative

splicing between two or more cell types or environments is observed.

Preferred splice-functional elements include sequences from the mammalian (e.g. rodent) B-tropomycin gene, from the mammalian fibroblast growth factor receptor-1 gene, from the gene for the γ subunit in mitochondrial ATP synthase and the mammalian (e.g. human) non-muscle myosin heavy chain B gene. The particular regions within these genes which may advantageously be used are described in the Examples.

The identification of further splice elements suitable for use in the invention can be carried out by methodology known in the art; by way of example methodology see Cote *et al* (1997 J Biol. Chem. 272:1054-1060), Balvay *et al* (1992 NAR 20: 3987-3992) and Guo and Helfman 1993 (NAR 21:4762-4768).

"Nucleotide sequences of interest" NOI's according to the invention can be any chosen sequence encoding an mRNA or protein of which expression or translation in a specific cell type or environment is desired. The product of the sequence of interest maybe a spliced mRNA transcript, e.g. for use in known anti-sense or ribozyme methodologies, or a protein. Proteins encoded by nucleotides of interest according to the invention will generally be classed as therapeutic proteins. For instance, such proteins might be toxic to a particular cell type, or activate a pro-drug. "Therapeutic proteins" include proteins which are themselves therapeutically active or have an indirect therapeutic activity, e.g. by enzyme action to convert a pro-drug into its active form. Reporter or marker proteins are also envisaged for testing the methods of the invention and imaging and/or identifying target cells. The NOI will typically be exogenous to the target cell-

population, e.g. encoding a drug or pro-drug, or exogenous in the sense that the equivalent naturally occurring gene is different, for example in classic gene therapy where a fully functional copy of a gene is introduced to overcome problems associated with a defective naturally occurring gene.

Specific cell types or environments according to the invention are those in which an alternative mechanism relative to other cells or cell environments can be elucidated and can include, but are not limited to those infected with viral or other infectious agents, benign or malignant neoplasms, or components of the immune system which are involved in autoimmune disease or tissue rejection. The means provided by the invention may be used to target cells and/or environments in any eukaryotic organism wherein alternative splicing is exhibited between cell populations and expression of a NOI in a given cell-population(s) is desired.

Thus, the present invention provides a novel method of targeting expression of a therapeutic gene to a particular cell population. For example malignant rather than normal cells of a given type or one or more cell types within a given tissue, organ or part of the body, e.g. skeletal as opposed to smooth muscle.

In a preferred embodiment of the present invention, the NOI, e.g. the therapeutic gene, is not expressed as a fusion product with one or more exons or partial exons from the splice-functional nucleotide sequences (elements). Fusion genes may be problematic as they will typically result in protein products with lower activity and may be localised to an inappropriate area of the cell by upstream endogenous sequences within the fusion gene.

According to the present invention therefore, the NOI is preferably not expressed as a fusion protein with any other regions of the genetic construct introduced into the target cells. Although the NOI may, in certain circumstances itself encode a fusion protein e.g. heavy and light chains of an antibody. It does not form a fusion protein with part of the (typically endogenous) splice-functional elements. The NOI ORF in a target cell is the primary ORF, the first (possibly only) ORF available for translation, it is not proposed that an mRNA AUG start codon from part of the splice-functional or surrounding elements is utilised to translate the NOI as a fusion protein therewith.

In a particularly preferred embodiment of the present invention, to ensure that translation begins from the correct start codon (ATG) in the target cells, upstream start codons are manipulated. Conveniently, upstream start codons are mutated e.g. ATG to ACG by techniques for site-directed mutagenesis which are well known in the art. In this way, the first available ORF of the mature mRNA in the target cells will be for the therapeutic gene. This also means that fusion proteins with exons or partial exons from the splice-functional elements may be avoided.

Preferably, when a region of the whole genetic construct is excluded by a splicing reaction in the target cells, one or more ORFs are retained in the region which forms part of the mature mRNA is non-target cells (but is excluded in target cells) which is upstream of the therapeutic gene. So, in non-target cells, the first available ORF belongs to a gene other than the therapeutic gene which it is desired to express in the target cells. Translation in the normal 5' cap dependent manner means that while in non-target cells the mRNA may contain the therapeutic gene it will not be translated or

it will be translated in small amounts. A system may of course be provided where two therapeutic genes are incorporated into the transcript, one for expression in one population of cells, the other for expression in a different population of cells, as exemplified by the luciferase and EGFP genes of Figure 1 and Example 1.

When the splicing mechanism means that in the target cell population the mRNA contains all the regions of the mRNA in non-target cells and some additional sequences, i.e. there is exon inclusion in target cells, the mRNA in target cells will preferably have just one ORF, i.e. for the therapeutic gene of interest and in non-targeted cells the mature mRNA will preferably have no ORFs.

The present invention utilises the cell's *cis*-splicing mechanism such that exons from the same pre mRNA molecule are used to produce a mature mRNA splice product. The use of trans-splicing techniques in therapy has been described (US 601348) but this involves formation of an mRNA duplex between a target mRNA molecule which must be specific to the cells of interest and an administered therapeutic mRNA. According to the present invention on the other hand, the normal *cis*-splicing mechanisms are utilised, targeting being achieved through the different effects of splice-functional elements in different cell populations.

The present invention utilises alternatively spliced molecules (ASMs) having the characteristics of the genetic constructs described above.

The alternatively spliced molecules (ASMs) of the invention can be RNA, DNA, or other nucleotide analogue molecules such as peptide nucleic acids (PNA). By "alternatively spliced molecules" is meant molecules that give rise to or encode a transcript which is

spliced differently in one cell type and/or environment relative to another. The differential, alternative or abherrant *cis*-splicing reaction provides an active desired molecule which can be expressed or translated in the target cells. The expression product of the desired mRNA may be a protein of therapeutic value to the cell, or a toxin which kills the targeted cells. Alternatively, the desired mRNA alternatively spliced in the target cell may itself perform a therapeutic function, for instance via, for example, known ribozyme or anti-sense technology. In another embodiment, the expression product derived from the alternative *cis*-splicing is secreted by the cell. In further embodiment of the invention multiple distinct ASMs may be used in combination to achieve a desired effect, for example, via the expression of multiple NOIs or targeting multiple cell types and/or environments with a single NOI.

A nucleotide sequence encoding a translatable protein capable of producing an effect, such as cell death, or alternatively, one that restores a missing function or acts as a marker, may be incorporated in the ASMs of the invention as the NOI. For example, the nucleotide sequence can include those sequences encoding gene products missing or altered in genetic diseases where the affected cell types display or are found e.g., upon assay to display, differential splicing relative to other cells. Alternatively, the nucleotide sequences can encode marker proteins or peptides which may be used to identify and/or image cells.

Genes or NOI for killing cells (encoding cytodestructive agents) may be simple toxins or other genes which provide some function which enhances the susceptibility of the cells to subsequent treatments, such as radiation or chemotherapy e.g. an enzyme capable of cleaving a

pro-drug. Toxins such as diphtheria toxin (Dl), ricin, Pseudomonas toxin, shiga toxin, exotoxin A, cholera toxin and other protein or peptide toxins for which a coding sequence is known in the art can be used in the invention.

In one embodiment, the toxin is active within the cell in which is expressed, but unable to leave the target cell to affect other cells. Diphtheria toxin is an example of such a toxin. Native DT is made up of an A subunit and a B subunit. Diphtheria toxin subunit A contains the enzymatic toxin activity and will function if expressed or delivered into human cells. The B subunit is required for transmembrane movement into human cells. Subunit A can not enter intact cells by itself. Alone, subunit A has very low toxicity because it cannot cross the lipid bilayer of a cell membrane without the B subunit. Donovan, I. M. Simon and M. Montal, J. Biol. Chem 260, 8817-8823 (1985.

In an alternative embodiment the expressed protein encoded by the ASM may be secreted from affected cells to the surrounding tissues.

The nucleic acid molecules and constructs of the invention can be RNA or DNA or derivatives or modified versions thereof, single-stranded or double-stranded. By nucleic acid is meant a ASM molecule or a nucleic acid molecule encoding a ASM molecule, whether composed of deoxyribonucleotides or ribonucleosides, and whether composed of phosphodiester linkages or modified linkages. The term nucleic acid also specifically includes nucleic acids composed of bases other than the five biologically occurring bases (adenine, guanine, thymine, cytosine and uracil).

The RNA and DNA molecules of the invention can be

prepared by any method known in the art for the synthesis of DNA and RNA molecules. For example, the nucleic acids may be chemically synthesised using commercially available reagents and synthesisers by methods that are well known in the art (see, e.g., Gait, 1985, Oligonucleotide Synthesis: A Practical Approach, IRL Press, Oxford, England). Alternatively, RNA molecules can be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the RNA molecule. Such DNA sequences can be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or 5P6 polymerase promoters. RNAs may be produced in high yield via *in vitro* transcription using plasmids as known in the art. In addition, RNA amplification methods such as Q-β amplification can be utilised to produce RNAs.

The nucleic acid molecules can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridisation, transport into the cell, etc. For example, modification of a ASM to reduce the overall charge can enhance the cellular uptake of the molecule. In addition modifications can be made to reduce susceptibility to nuclease degradation. The nucleic acid molecules may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane or the blood-brain barrier. Various other well-known modifications to the DNA molecules can be introduced as a means of increasing intracellular stability and half-life.

The nucleic acids may be purified by any suitable means, as are well known in the art. For example, the nucleic acids can be purified by reverse phase chromatography or gel electrophoresis.

In instances where a nucleic acid molecule encoding a ASM is utilised, cloning techniques known in the art may be used for cloning of the nucleic acid molecule into an expression vector. Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), 1993, Current Protocols in Molecular Biology, John Wiley & Sons, New York; and Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, New York.

In a further embodiment, the present invention provides a vector comprising a genetic construct as defined herein. In addition, the present invention provides a vector comprising the splice-functional elements as described herein with a site suitable for NOI insertion wherein the splice functional elements are selected and arranged in such a way that the first ORF available for translation in a target cell population is for the NOI while in non-target cells the NOI ORF is absent or not the primary ORF.

The DNA encoding the ASM of interest may be recombinantly engineered into a variety of host vector systems that also provide for replication of the DNA in large scale and contain the necessary elements for directing the transcription of the ASM. For example, a vector can be introduced *in vivo* such that it is taken up by a cell and directs the transcription of the ASM molecule. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art.

Vectors encoding the ASM of interest can be plasmid, viral, or others known in the art, that are used for replication and expression in mammalian cells.

Expression of the sequence encoding the ASM can be regulated by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Benoist, C. and Chambon, P. 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. NatI. Acad. Sci. U.S.A. 78:14411445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, Nature 296:39-42), the viral CMV promoter, the human chorionic gonadotropin-β promoter (Hollenberg et al., 1994, Mol. Cell. Endocrinology 106:111-119), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced into the desired tissue site.

Alternatively, viral vectors can be used which are capable of infecting the desired target cell. Vectors according to the invention may be based upon vectors known in the art, for example plasmid-based vectors, retroviral based vectors, lentiviral-based vectors, adenoviral based vectors and combinations of such vectors. Viral vectors with transient expression, such as adenovirus and adeno-associated virus, which do not integrate into the genome can also be used in the invention.

Methods of generating genetic constructs and vectors having the characteristics described herein constitute further aspects of the present invention; suitable methods are well known in the art and described in the Examples herein.

Additional features can be added to the ASM molecule of

the invention, or the vector in which it is carried, in addition to the desired gene or NOI encoding the desired mRNA, such as polyadenylation signals, additional splice sites, stop codons, or protective groups to modulate the stability of the molecule. As discussed above the ATG start codons upstream of the NOI may be mutated. Further elements such as a 3' hairpin structure, circularised RNA, nucleotide base modification, or synthetic analogues) can be incorporated into constructs to promote or facilitate nuclear localisation and spliceosomal incorporation, and intra-cellular stability.

The method of the present invention does not require delivery to only the targeted cells. The precursor molecule to the desired molecule can be delivered and taken up by multiple cell types, for example, all cells in the organism, but the desired mRNA is only functionally spliced *in vivo* in the specific target cells to express the NOI. The specificity of this technique relies on the unique (restricted) transcription of the desired RNA encoding the NOI in the target cells. Preferably, the non-targeted, e.g. non-diseased cells will not translate, or will at most translate only minimally the desired gene or NOI. Therefore, selective expression of the desired molecule will not take place in such non-targeted cells, or will only take place to a very minimal extent, i.e. at therapeutically or physiologically insignificant levels.

The ASMs of the invention can be used in methods designed to produce a desired RNA in a target cell. The methods of the present invention comprise delivering to the target cell a ASM which may be in any form used by one skilled in the art, for example, an RNA molecule, or a DNA vector which is transcribed into a RNA molecule.

Therefore the invention provides in a further embodiment a method of targeting gene expression to a cell-population of interest comprising the steps of:

i) delivering to the nucleus of cells a precursor molecule, which may be of any form used by one skilled in the art, an alternatively spliced molecule (or a DNA vector which may be transcribed into RNA or a synthetic analogue) which molecule contains, splicing elements, a desired NOI, and may contain polyadenylation signals, enhancers, or other modulating sequence elements;
ii) *cis*-splicing of the alternatively spliced molecule to create a desired mRNA encoding the NOI; and
iii) expression (translation) of the *cis*-spliced desired mRNA encoding the NOI within the target cell(s).

The ASM can be administered to cells by any delivery procedure, for example, virally mediated, electroporation, micro injection, calcium phosphate mediated transfection, liposomes, cytofectins, or directly. The ASM molecule will be administered in amounts which are effective to produce the desired amount of the expression product of the NOI molecule itself. The exact amount administered will vary depending upon the details of the delivery system. The effective amount may also vary depending upon whether the expression product of the NOI provides a missing function (such as in therapy of genetic disease), cell death (such as in therapy of cancer), or cell regulation (which may be used with different types of diseases). The effective amount may range from 0.001 pico g, or even less, to 1.0 g nucleic acid/kg body weight of the patient.

Various delivery systems are known and can be used to transfer the compositions of the invention into cells, e.g. encapsulation in liposomes, microparticles,

microcapsules, recombinant cells capable of expressing the composition, receptor-mediated endocytosis, construction of a nucleic acid as part of a retroviral or other vector, injection of DNA, electroporation, calcium phosphate mediated transfection, etc.

The present invention can use any known delivery system developed for other gene treatment or anti-sense methods. The ASM encoding RNA can be made, packaged, and tested for cellular incorporation if required. The person skilled in the art will realise that ASMs according to the invention can be purified and tested by sequencing or by ability to *cis*-splice correctly to express the NOI in an *in vitro* system.

Liposomes, electroporation, and cytofections are methods of directly introducing RNA into cells. They are widely used in anti-sense RNA delivery protocols. Naked or packaged DNA is another possible means of delivery. Viral delivery systems may also be used as described herein. Nucleic acid polymers can also be delivered by attaching them to the empty shells of replication incompetent viruses. Cook, D. R. et al. Cancer Biother. 9(2), 131-141 (1994).

Delivery of the nucleic acid into a host cell may be either direct, in which case the host is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, host cells are first transformed with the nucleic acid *in vitro*, then transplanted into the host. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene delivery.

The final therapeutic construct may have to be delivered from the external cell membrane and transit into the nucleus to be incorporated into the spliceosome of the

targeted pre-mRNA, or they may be produced by the cell itself from a precursor molecule (DNA vector, DNA virus, etc.).

In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the ASM. This can be accomplished by any of numerous methods known in the art, e.g., by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g. by infection using a defective or attenuated retroviral or other viral vector (see U.S. Pat. No. 4.980286), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432).

In a specific embodiment, a viral vector that contains the ASM can be used. For example, a retroviral vector can be utilised that has been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA (see Miller et al., 1993, Meth. Enzymol. 217:581-599). Alternatively, adenoviral or adeno-associated viral vectors can be used for gene delivery to cells or tissues. (See, Kozarsky and Wilson, 1993, Current Opinion in Genetics and Development 3:499-503 for a review of adenovirus-based gene delivery).

Another approach to gene delivery into a cell involves

transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. The resulting recombinant cells can be delivered to a host by various methods known in the art. In a preferred embodiment, the cell used for gene delivery is autologous to the host cell.

The present invention also provides pharmaceutical compositions comprising an effective amount of a ASM or a nucleic acid encoding a ASM, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognised pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Examples of suitable pharmaceutical carriers are described in "Remington s Pharmaceutical sciences" by E. W. Martin.

In specific embodiments, pharmaceutical compositions are administered in diseases in which it has been shown that diseased cells exhibit an alternative splicing mechanism as compared to non-diseased cells, in particular: (1) in diseases or disorders involving an absence or decreased (relative to normal or desired) level of an endogenous protein or function, for example, in hosts where the protein is lacking, genetically defective, biologically inactive or underactive, or under expressed; or (2) in diseases or disorders wherein, *in vitro* or *in vivo*, assays indicate the utility of ASMs that inhibit the

function of a particular protein.

The activity of the protein encoded for by the mRNA resulting from the *cis*-splicing of the ASM can be readily detected, e.g., by obtaining a host tissue sample (e.g., from biopsy tissue) and assaying it *in vitro* for mRNA or protein levels, structure and/or activity of the expressed NOI. Many methods standard in the art can be thus employed, including but not limited to immunoassays to detect and/or visualise the protein encoded for by the mRNA (e.g., Western blot, immunoprecipitation followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridisation assays to detect formation of mRNA expression by detecting and/or visualising the presence of ASM mRNA (e.g., Northern assays, dot blots, in situ hybridisation, and Reverse-Transcription PCR, etc.), etc.

In a specific embodiment, it may be desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment. This may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. Other control release drug delivery systems, such as nanoparticles, matrices such as controlled-release polymers, hydrogels.

The ASM will be administered in amounts which are effective to produce the desired effect in the targeted cell. Effective dosages of the ASMs can be determined through procedures well known to those in the art which

address such parameters as biological half-life, bioavailability and toxicity. The amount of the composition of the invention which will be effective will depend on the nature of the disease or disorder being treated, and can be determined by standard clinical techniques. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges.

The present invention provides a new method of targeting in gene therapy techniques and thus in a further aspect the present invention provides constructs, ASM's and vectors as described herein for use in therapy. More particularly, these moieties are suitable for use in the treatment of diseases which may be ameliorated by gene therapy e.g. diabetes or sickle cell anaemia or cancer.

The present invention also provides the use of the constructs, ASM's and vectors as described herein in the manufacture of medicament whose physiological effect is restricted to a target cell population and which treats a disease characterised in that it is responsive to expression of an exogenously administered gene. The disease may result from the lack of normal expression of a given gene which may be added as an NOI e.g. proinsulin or in the case of cancer it may be desirable for the NOI to be a toxin which kills only the malignant cells. The NOI may express a dominant negative protein, used in therapy to interfere with the activity in the cell of an endogenously expressed protein, e.g. by forming an inactive heterodimer therewith.

In the field of cancer research there are now an increasing number of examples whereby specific RNA splicing is abherrant/alternative. Indeed it is often the products of such splicing that are implicated in some way in the progression of oncogenesis.

Such examples include that which occurs in the RNA message for TSG101 (Lin *et al* 1998 Br J. Haematol. 1998 102: 753-8: Lee and Feinberg 1997 Cancer Res. 57; 3131-4), cathepsin B (Mehtani *et al* 1998 J Biol Chem. 273 13236-44), EGF receptor (e.g. see Yamazaki *et al* J. Natl. Cancer Inst. 1998 90: 581-7; Sugawa *et al* 1990 PNAS: 87:8602-6), CD44 (see Kittl *et al* 1997 Exp. Clin. Immunogenet 14; 264-72; Sherman *et al* 1997 Cancer Res 57; 4889-97), NER (Saito *et al* Oncogene 1997 14: 617-21), PDGF receptor (Mosselman *et al* 1996 PNAS 93; 2884-8), c-*kit* (Takaoka *et al* 1997 Cancer Lett. 115; 257-61), FGF receptor (Luqmani *et al* 1995 64; 274-9), Glucocorticoid receptors (Moalli and Rosen 1994 Leuk Lymphoma 15: 363-74), and p53 (Nakai *et al* 1994 Br. J. Haematol.).

Often in these and other examples, changes in splicing patterns of certain genes is not due to primary mutations found within the gene. These examples suggest that abherrant/alternative splicing has some relationship to the development/progression of a variety of cancers. Indeed the correlation between the presence of certain abherrant/alternative spliced transcripts and emerging or residual cancers is such that the presence of such transcripts can be used as molecular markers for diagnosing disorders (see Taniguchi *et al* US patent 5643729)

Often for such abherrant/alternative splicing in disease, the *precise* mechanism by which exons are skipped remains to be elucidated. Nonetheless these events are selected for and be they either cause or effect of disease progression such abherrant splicing is a useful targeting method with which to ensure NOI targeting.

Included within the scope of the invention is the

application of the invention *to in-vivo* and *ex-vivo* therapeutic regimes as currently undertaken in gene therapy and gene delivery protocols.

The present invention also provides a pharmaceutical pack or kit comprising one or more compartments containing a construct, ASM or vector as described herein optionally associated with such compartment(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human or animal administration.

In further preferred aspects, the present invention also provides:

The use of rat B-tropomycin exon 5 splice donor (and surrounding functional sequence) and/or the use of rat B-tropomycin exon 7 splice acceptor (and surrounding functional sequence) to confer cell specific splicing events that limit optimal NOI translation/expression to either skeletal muscle or non-muscle cells. The NOI delivery vector maybe of any type: Plasmid-based vectors, retroviral based vectors, lentiviral-based vectors, adenoviral based vectors and combinations of such are example such NOI delivery vectors. The promoter and polyadenylation signals used to mediate the transcription of the RNA transcripts as described can be of any origin.

The use of B-tropomycin exon 5 splice donor (and surrounding splice-functional sequence) from any species and/or the use of B-tropomycin exon 7 splice acceptor (and surrounding splice-functional sequence) from any species to confer cell specific splicing events that limit optimal NOI translation/expression to either

muscle or non-muscle cells. The NOI delivery vector maybe of any type: Plasmid-based vectors, retroviral based vectors, lentiviral-based vectors, adenoviral based vectors and combinations of such are examples. The promoter and polyadenylation signals used to mediate the transcription of the RNA transcripts as described can be of any origin.

The use of human FGFr α-exon (and surrounding splice-functional sequence) to confer cell specific splicing events that limit optimal NOI translation/expression to glioblastoma related/derived cells. The NOI delivery vector maybe of any type: Plasmid-based vectors, retroviral based vectors, lentiviral-based vectors, adenoviral based vectors and combinations of such are example. The promoter and polyadenylation signals used to mediate the transcription of the RNA transcripts as described can be of any origin.

The use of FGFr α-exon (and surrounding splice-functional sequence) from any species to confer cell specific splicing events that limit optimal NOI translation/expression to glioblastoma related/derived cells. The NOI delivery vector maybe of any type: Plasmid-based vectors, retroviral based vectors, lentiviral-based vectors, adenoviral based vectors and combinations of such are examples. The promoter and polyadenylation signals used to mediate the transcription of the RNA transcripts as described can be of any origin.

Preferably, in the above described uses, removal/ mutation of any unrequired upstream ATG translation initiation sites is used to ensure translation initiation at the desired ATG such that translation/ expression of the NOI is optimal.

Moreover, the use of other splice sites and surrounding splice-functional elements from any origin (either natural or synthetic) in a manner and configuration similar to that described for examples 1 and 2 of the text to confer cell type specific splicing such that NOI optimal translational/expression is restricted to specific cell types and/or environment.

In addition, the methods, uses and constructs described herein may be used to in conjunction with other methodologies, protocols and regimes, particularly to confer an even greater degree of selectivity of expression, e.g. by the use of tissue specific promoters.

The various methods and uses described herein have *in vivo* and *ex vivo* applicability.

The invention will now be described in more detail with reference to the following non-limiting Examples. The Examples illustrate the use of abherrant or alternative splicing in the restriction of gene expression to certain environments and cell types. They aid in understanding the current invention and are therefore not limitations of the invention to the scenarios of the examples. In these Examples:

Figure 1(A) shows a diagram of the pCI vector incorporating splice donor and acceptor sequences from rat B-tropomysin and two reporter genes, (B) shows in schematic form how unspliced transcripts are produced in non-skeletal muscle cells and spliced transcripts in skeletal muscle cells.

Figure 2 shows in schematic form (A) the vector pFGR-17 which incorporates a marker gene and, (B) how EGFP expression fails to occur in healthy astrocytes and (C)

how EGFP expression is confined to glioblastoma cells.

Figure 3(A) shows in schematic form the F1γ mini-gene expression muscle and (B) the vector JC-CAGGS-F1γ incorporating it, (C) shows the positioning of the marker genes within the vector (D) the mature mRNA transcripts in non-muscle (luciferase + EGFP) and muscle (only EGFP) cells.

Figure 4(A) shows in schematic form the genomic sequence for part of human non-muscle myosin heavy chain B, (B) a mini-gene expression cassette incorporating it, (C) this cassette incorporating the NOI and (D) the alternatively spliced forms of the mature transcripts the longer transcript predominating in neural cells.

Figure 5 shows in diagrammatic form how cell-specific splicing elements may be identified using genomic DNA libraries.

EXAMPLES

EXAMPLE 1: Use of B-tropomysin exon7/upstream intron regulatory sequences to restrict gene expression to skeletal muscle cells

Outlined in Figure 1A is a representative diagram of a vector (in this instance pCI (Promega) based) in which the pCI intron related sequences are replaced with the splice donor/flanking sequence (50 base pairs either side of the splice donor) from exon 5 of rat B-tropomysin and the splice acceptor/flanking sequence (from upstream of the branch point through 25 nucleotides downstream of exon 7) of rat B-tropomycin. Between these two sequences is inserted a luciferase gene taken from pGL3 (Promega) with its ORF orientated as shown. Downstream of all these sequences is located a EGFP (enhanced green fluorescent protein) ORF

(Clontech) as shown. The poly A signal is as found downstream of the multiple cloning site of pCI. To ensure translation begins from the desired ATG (of either luciferase or EGFP) then other, upstream ATG sequences present in the primary transcript are mutated (in this instance by ATG to ACG mutagenesis) This vector can be made by established methods by those skilled in the art.

Because of the blockage in splicing conferred by the exon 7 sequence in non-muscle cells, the resulting vector of Figure 1 will now only splice in skeletal muscle cells. As a consequence of this design, in skeletal muscle cells there is observed an increase in EGFP expression and a decrease in luciferase expression. In non-skeletal muscle there is instead an observed increase in luciferase expression and a decrease in EGFP expression. Consequently, in this example EGFP has skeletal muscle specific high expression and luciferase has non-muscle specific high expression. Of note is that although when unspliced, EGFP is still present within the RNA transcript, without the use of IRES sequences, ribosome translation will only lead to luciferase translation in a 5'-cap-dependant manner.

This example therefore demonstrates the potential for the use of *cis* elements to limit expression of genes to certain environments. Although in this instance marker/reporter genes are used, these examples are not limiting. Other NOI of interest or of therapeutic value could instead be used.

EXAMPLE 2: Use of *cis* elements involved in FGFr exon skipping to target gene expression to glioblastoma cells

Progression of astrocytes from benign to a malignant phenotype is accompanied by a change in the RNA

processing of fibroblast growth factor receptor-1 gene. The level of high affinity form of the FGFr is dramatically increased due to α-exon skipping during RNA splicing. The *cis* regulatory sequence requirements to confer this exon skipping have been elucidated (Cote *et al* 1997 J Biol. Chem. 272:1054-1060) by use of a vector construct pFGR17 (Cote *et al* 1997 *i.b.i.d.)*. This was originally made by cloning the α-exon, γ-exon and flanking sequence into intron 1 of the metallothionein (MT) gene. Consequently transcripts produced from this vector, when spliced, have hMT exon1, α-exon, and hMT exon 2 of the mRNA joined. However if the inserted FGFr sequence induces exon skipping then α-exon sequence would be excluded in the final mRNA message. Consequent stepwise deletion of the inserted sequence was able to demonstrate that a 375 nucleotides sequence inclusive and flanking that of the α-exon, was able to confer the exon skipping phenotype, provided the overall insert size was maintained. These sequences had splice-functional *cis* elements involved someway in altering splicing in glioblastoma cells.

The pFGR-17 vector used to map the *cis* elements involved in FGFr exon skipping is the starting construct of this aspect of the invention- namely to generate glioblastoma specific NOI translation/expression by conditional splicing. Into the vector is first cloned a NOI, in this case the EGFP marker. This NOI is cloned downstream of the splice acceptor related sequences of hMT-2 exon of pFGR-17. See Figure 2 for schematic diagram of vector design. Next, to ensure that upon splicing, translation starts from EGFP ATG and not upstream ATGs found either within hMT-1 and hMT-2 sequence, such upstream ATGs are disabled (by ATG to ATTG insertion mutagenesis in this instance). This vector can be made by established methods by those skilled in the art.

In this instance the presence of ORFs (open reading frame) within the α-exon/flanking regions ensures that unless exon skipping occurs, EGFP is never translated. However in glioblastomas, α-exon induced skipping leads to upstream ORF removal and subsequently EGFP is then the first ORF of the message and as such is translated. Consequently only during exon skipping (as does occur in glioblastomas) is EGFP expressed optimally (Figure 2C), expression fails to occur in healthy astrocytes (Figure 2B). Although in this instance a marker gene is used, other NOI of interest or of therapeutic value could instead be used to replace EGFP, using established methods by those skilled in the art.

EXAMPLE 3:

It has been shown previously that exon 9 of the mitochondrial ATP synthase γ subunit is selectively excluded in both skeletal muscle and heart tissue (Matsuda et al (1993) FEBS Lett 325; 281-284, Matsuda et al (1993) J.Biol. Chem. 268; 24950-24958 and Ichida et al (1994) 273; 8492-8501). Subsequent to this initial observation it also been demonstrated that there exists a *cis*-acting element responsible for this exon 9 exclusion in such tissue (Masaru et al (2000) J. Biol. Chem. 275; 15992-16001). As with other such elements, and as would be expected, this element is located within the vicinity of the exons and introns in question. On this particular occasion it was shown to be located within exon 9 and was identified by the normal methodology and with the aid of a typical mini-gene expression cassette.

This observed cell-type specific mRNA splicing is exploited to generate a muscle specific NOI expression vector. As with previous such embodiments this is a non-limiting example outlined only such to aid in the

understanding of the invention. The starting sequence is a F1γ mini-gene expression cassette as outlined in Figure 3A. This consists of the exon in question as well as the neighbouring exonic and intronic sequence. The starting vector for this embodiment is named JC-CAGGS-F1γ (see Masura *IBID* for construction and Figure 3B for schematic representation). This vector consists of a pol II based promoter driving the F1γ mini-gene cassette of Figure 3A. Specifically it consists of the genomic sequence from Exon 8 through to Exon 10 (thus also including intron 8, exon 9 and intron 9). In this vector, mRNA polyadenylation is achieved by way of a Rabbit β globin poly A signal located downstream of the mini-gene. This vector has previously been used to demonstrate exon 9 exclusion to be specific to heart and skeletal muscle tissue in transgenic mice. It is related to pcDEB- F1γ (not shown) which contains the same mini-gene and used previously to demonstrate the same exon 9 exclusion to also occur ex-vivo during differentiation of myoblast cells (Masura *IBID*).

In the embodiment of the invention now described, the JC-CAGGS-F1γ is so modified such as to obtain muscle-cell specific expression of a desired marker gene. This marker gene can be replaced, if preferred, by other NOIs as and when desired. For this embodiment the JC-CAGGS-F1γ vector is first modified by the insertion of a luciferase reporter gene downstream of the exon 9 located, *cis*-acting element involved in cell specific exon 9 exclusion (shaded in grey in Figure 3A, B and C). Subsequent to this insertion the second reporter gene (EGFP) is then inserted within exon 10 of the same mini-gene. Both such insertions are placed so as not to interfere with normal GT and AG splice site related signals. Finally all translation initiation signals (ATGs) located upstream of EGFP gene (except those found in the inserted Luciferase gene) are then mutated (by

ATG to TTG mutagenesis in this instance).

A schematic of this vector, named pMus-Spec is represented in Figure 3C and can be readily constructed by those skilled in the art. This vector now contains two ORFs. The first (luciferase) in exon 9 whilst the other (EFGP) in exon 10. As a consequence, if exon 9 is retained the luciferase gene is thus translated in the normal 5' cap dependant manner. However upon the exclusion of exon 9, as is observed in heart and skeletal muscle tissue, the luciferase ORF is thus also excluded. Consequently in such transcripts the EGFP ORF is now the first ORF available for translation. The pMus-Spec vector thus confers muscle specific gene expression on the EGFP ORF. See Figure 3 D for schematic outline of pMus-Spec mature mRNA transcripts. The longer luciferase expressing transcript will predominate in non-muscle cells whilst the shorter EGFP expressing transcript will do so in muscle cells.

EXAMPLE 4: Neural specific gene expression via *cis*-acting elements of human non-muscle myosin heavy chain B (NMHC-B) transcript

The transcript encoding human non-muscle myosin heavy chain B (NMHC-B) is adapted and re-engineered such to confer neural specific gene expression on an appropriate and desirable NOI. This is possible because the mRNA transcripts of the NMHC-B gene have previously been shown to contain a *cis*-acting element involved in the regulation of mRNA transcript splicing. As is typical, this *cis*-acting sequence is found located within the locality of the exons (termed N30 and R18) found alternatively spliced; both of which are only included in neural cell specific transcripts. In this instance, and with aid of a traditional mini-gene cassette format, the location of this element has been identified and is

found downstream of both exons (also see Figure 4A; description outlined in text below). Because both the N30 and R18 exons are only included for translation in neural cells, this *cis*-acting element has as a consequence been termed an intronic distal downstream enhancer (IDDE). It has been shown that when expressed in a mini-gene expression cassette, exon inclusion occurs 87.7 (+/-6.4) % of the time in the neural Y79 cell-line whilst only 2.1 (+/- 1) % of the time in the non-neural Hela cell line. More in depth analysis has subsequently shown that the polypyrimidine tract of the splice site preceding the N30 exon is sub-optimal and that the IDDE 'rescues' this tracts splicing function only in the neural cells (Guo and Kawamoto J. Biol. Chem. August 7; 2000 as Publication, manuscript M005597200).

The starting genomic sequence and subsequent mini-gene expression vector for this particular embodiment is as previously described (Guo and Kawamoto *IBID*). In brief this consists of the following: First the genomic sequence is outlined in Figure 4A and consists of exons E5, N30, R18 (the two neural specific exons) and E6. Also shown are the neighbouring intronic sequence (indicated as lines) and the IDDE (indicated as a black circle). For construction of the mini-gene expression cassette (see Figure 4B and Guo and Kawamoto *IBID*), the genomic sequence outlined in Figure 4B is inserted between the exons E2 and E3 of the rat preproinsulin gene (PPI; shown in stippled grey). Transcription of this mini-gene is driven by the Rous sarcoma virus long terminal repeat (RSVLTR). Once translated, the NMHC-B genomic fragment in this construct gives rise to two alternatively spliced products; those which include N30 and those which exclude it.

To aid in understanding the invention, this mini-gene

expression cassette can be converted into one that maintains the desirable NOI in only neural cells. This is achieved by first insertion of the desired NOI into the N30 exon as shown in Figure 4C. As a consequence this NOI will only be included in transcripts expressed in neural cells. If the NOI concerned also requires translation then all upstream exonic translation initiating ATGs must also first be mutated (in this instance by ATG to ATC alteration) such that the ATG of the NOI is thus used for initiation. Such an embodiment can be constructed by those skilled in the art. A schematic outline of this neural specific NOI expression cassette is shown in Figure 4C and the alternatively spliced forms of the mature transcripts outlined in Figure 4D. The longer, NOI containing transcript predominate in neural cells.

EXAMPLE 5:

Alternatively cell-type specific splicing elements can be identified by the construction of a eukaryotic expression vector containing an intron and suitable gene (for example EGFP) into which is inserted a library of random sequence elements flanking the splice sites. This resulting 'splicing-regulatory element library' could then be delivered to a desired cell type, under the desired conditions and a gene expression screen/cell purification undertaken; by for example FACS sorting. Resulting cells, with the desirable gene expression characteristics will have then have their contained vector's flanking sequences sequenced. By way of example, a gene expression screen could be developed whereby a selectable marker is placed within an intron and so as such is not expressed unless splicing is inhibited. Into this vector could be cloned a sheared genomic library, such that the library sequences flank the surrounding intron.

The screen would then be to transfect such a library to a desirable cell population, treat the cell population in the desirable way and then subsequently identify clones expressing the selectable marker by use of a the appropriate selection. Desirable treatments might include mitogen/oncogene stimulation of the cell population prior to selection, and then subsequent identification of clones that express the selectable marker (due to splicing inhibition) only when growth stimulated (growth stimulation occurs during oncogenesis).